Article can be accessed online at http://www.publishingindia.com

Granular Region-oriented Fuzzy-Rough based kNN Improvization for Activity Recognition Modeling

Vijay Borges*, Wilson Jeberson**

Abstract

Activity recognition is a complex task of the Human Computer Interaction (HCI) domain with ever-increasing research interest. Human activity recognition has been specially addressed by the advances in pattern recognition. k-Nearest Neighbors (kNN) is a non-parametric classifier from pattern recognition theory, that mimics human decision making by taking previous experiences into consideration for segregating unknown objects. A novel fuzzy-rough model, based on granular computing for improvization of the kNN classifier is proposed herewith. In this model, feature-wise fuzzy memberships are generated to fuzzify the feature space of the nearest neighbours of the test object. These neighbor's fuzzified feature space are then aggregated into granules, based on their class-belongingness. From these, lower and upper approximation granules are generated using rough set theory to classify the test object. It is shown experimentally that this model outperforms the traditional kNN by 16.43% and Fuzzy-kNN by 10.25%, in the human activity recognition domain. Another novelty is in the efficient use of the fuzzy similarity relations in class-dependent granulated feature space, and, fuzzy-rough lower/upper approximations in the hybridization of the kNN classifier.

Keywords: k-Nearest Neighbors, Human Activity Recognition, Smart Environments, Pervasive Computing, Fuzzy Rough Sets, Fuzzy-rough Granular Computing

1. Introduction

Pervasive Computing is a technique of unobtrusively gathering data from various devices/sensors, process information from various sources; to both control physical processes and interact with human beings. Environments having pervasive computing facilities are called as Smart Environments [1]. Due to complex human nature, activities of humans are difficult to recognize. This work focuses on Activity Recognition (AR), that would use the underlying smart environment to infer/reason the activities of living entities (mostly humans) in an integrated manner. AR has many applications right from, security surveillance for mitigating threat to life/property, activity recognition in smart homes, assistive living for the elderly, nursery care for children, smart hospitals, smart educational institutes etc. [1].

Human activities are complex, as there can be multiple ways to do the same activity; either concurrently/ interleaved or in a shared fashion. Human activity interactions could be redundant, incomplete, overlapping and many times missing/ambiguous. Thus vagueness and incompleteness are two primary hallmarks of human activities. Vagueness and incompleteness, is ever more pronounced in individuals effected with Dementia diseases [2]. It is these individuals effected with Dementia, whose activities have to be monitored, which is of primal importance to this research effort. The data generated by the smart environment corresponding to human

* Assistant Professor, Department of Information Technology and Engineering, Goa University, Goa, India.
E-mail: vijayborges@gmail.com

activity interactions could be redundant, incomplete, overlapping and many times missing/ambiguous [3]. The kNN classifier is an established classifier that allots a test object to a decision class most commonly seen among its k-closest neighbors. But, the kNN classifier suffers loss in its accuracy in the domain of human activity recognition due to issues just pointed. Fuzzy-Rough set theory [4] [5], is a natural computing paradigm which attempts to characterize vague/incomplete data/information in a human-like form. It models vague (not clearly understood or expressed) and incomplete information, with a degree of certain/possible belongingness to a concept (decision class). Extending fuzzy set theory to kNN classifier, allows partial membership of an object to different decision classes and also shows measures of the importance (strength) of those memberships to it. Similarly, rough set theory extends the kNN classifier to handle approximations of concepts when faced with incomplete information. Vagueness and incompleteness are two primary hallmarks of human activities.

Information Granules [4], are collection of entities that are accumulated together due to their similarity, indistinguishability, cohesion, etc. Fuzzy memberships are built from the k-nearest objects, producing fuzzified feature space. The fuzzified k-nearest feature space is granulated in a class-dependent fashion to get the lower/upper approximated rough neighbors. This granulated rough neighbors hold better class distinguishing information, which is exploited in activity recognition. In this work we have shown experimentally that using the hybridized kNN with granulated fuzzy-rough extensions, the complex task of human activity recognition gives better classification accuracy over the traditional kNN classifier. The novelty of this work is that the fuzzy memberships are learnt from the k-nearest neighbors to the test object. This eliminated the task of relying on domain experts to come with fuzzy memberships that would be essential in getting discriminating Fuzzy Rules [6]. Another novelty is in the efficient use of fuzzy similarity relations in class-dependent granulated feature space, and, fuzzy-rough lower/upper approximations in hybridization of the kNN classifier, which have shown better performance over traditional kNN adapted as per the Parzen windows technique [7]; which considers objects within a fixed distance.

The rest of the paper is organized as follows. Section 2, introduces the necessary theoretical background of fuzzy sets, rough sets and touches upon granular computing. In Section 3, the proposed hybridized Fuzzy-Rough kNN is presented: various fuzzy similarity relations, fuzzy implicators and mechanisms to calculate the upper/lower approximation over the entire feature space to each concept is shown. Experimentation is performed on the proposed hybridized Fuzzy-Rough kNN using real-life dataset and is reported in Section 4. The paper concludes in Section 5 with future directions.

2. Theoretical Background

Granular Computing [4], is a framework for representing information in some form of aggregation (grouping a number of individual entities) and their processing. These information granules are very natural and characterize human thinking and reasoning. Granules of time, could be seconds, day, weeks, month, years; in natural language, word granules are essential entities for communication between humans. The grouping of these entities into an information granule is due to their similarity, indistinguishability, cohesion, etc; generally originating at numeric level. Granular Computing, is knowledge-inclined processing and employs various perception-based soft computing techniques like artificial neural networks, fuzzy set theory, rough set theory, fractal geometry, evolutionary algorithms. There are a number of formal well investigated frameworks for building information granules like, Set Theory and Interval Analysis, Fuzzy Sets, Rough Sets, Probabilistic Sets, etc. Information Granule A, in some space X, is a mapping $A : X \rightarrow G(X)$; where G, could be any one of the above stated formal frameworks. In this work, information granule would be limited to using Fuzzy-Rough set theory framework. Some formal theoretical background about fuzzy set theory and rough set theory would be elaborated in the following subsections.

2.1. Fuzzy Set Theory

In traditional set theory, an object is said to belong to a subset in a given universe of objects with crisp boundaries: which means that the element either belongs (part of the subset) or does not belong (not part of the subset). Due to humans psychoanalytical classification mechanism, traditional set theory falls short in capturing humans vague decision boundaries. Fuzzy set theory, allows to associate a degree of membership to the belongingness of an object to a subset. This degree of membership allows vague decision boundaries to crop in, as to capture humans psychoanalytical classification mechanism. For example, a set of physical characteristics of humans like, *short*, *obese*; or human activities like, *sleeping*, *toileting*, *eating*, cannot be aptly portrayed with crisp boundaries but exhibit vague boundaries. To capture these vague boundaries for decision-making, fuzzy set theory plays a pivotal role.

Given a universe of objects U, and a class C. An object $x \in U$ (also called *input pattern, feature vector)*, is said to belong to C, depending on the characteristic function $\mu_C : U \to [0,1]$. This characteristic function defines Fuzzy Sets; which associates to every object x an associated degree of membership to the class C, as $\mu_C(x)$, and formally shown as $\{x, \mu_C(x)\}$. Fuzzy set can degenerate to a crisp set if the characteristic function is defined as $\mu_C : U \to \{0, 1\}$, or $\mu_C(x) = 1$, if $x \in C$ and $\mu_C(x) = 0$, otherwise. Thus fuzzy set theory is a natural extension to the traditional set theory. The advantages of fuzzy set theory is of paramount importance to the domain of activity recognition; where, *human interactions* in smart environments have to be clearly classified to recognizable *activities.* Since human interactions are generally vague, overlapping and has no well defined boundaries; fuzzy technique of associating *human interactions* to various degrees of memberships to the *activity classes*, is very useful.

2.1.1. Fuzzification of Feature Space By Learning from Data

The criticality of using fuzzy set theory lies in the ability of coming with truly discriminating Fuzzy Membership Function (FMF), so that the input feature space can be correctly classified to appropriate output decisions/classes. It depends on domain expert to come with these discriminating FMF's. There are challenges though to comeup with these FMF's, like: (a) FMF's are problem dependent, there are hardly any *"one-fit-all"* FMF's for a class of problems; (b) FMF's may suffer due to presence of absurd/noisy input data; (c) due to the dynamic nature of the problem space, static FMF's from domain expert may be difficult to obtain. The above stated challenges become more pronounced in the domain of human activity recognition in smart environments; where human activities of daily living can be done by two or more individuals in parallel, concurrently, incomplete or interleaved fashion [3]. In addition to that the underlying firmware of the smart environment could play its dubious role in generating incomplete, noisy and redundant data of interactions of humans in the environment. In this work, we propose a novel model of building these FMF's by learning from the input data itself, and thus eliminating the need of a domain expert. The proposed method can also be extended to incorporate linguistic/contextual information coming from domain experts; if required.

Let $X = \{\underline{x}_1, \underline{x}_2, .., \underline{x}_N\}$ where $\underline{x}_i \in \mathrm{R}^p$, be a set of input patterns (feature vectors) [human activity interactions in smart environment] and $A = \{a_1, a_2, .., a_M\}$, be a set of pre-defined output classes [human activities/activity classes/decision classes]. The requirement is to generate fuzzy rules using a mapping, $f{:}\{\underline{x}_i, .., \underline{x}_j\} \to a_k$. Let $\underline{x}_i[k]$'s, minimum-maximum interval range be $[\underline{x}_i[k]^-, \underline{x}_i[k]^+]$. Using fuzzy set theoretic technique [6][5], $\underline{x}_i[k]$'s, interval range is divided into three fuzzy regions, *low, medium, high.* The $\underline{x}_i[k]$, is then assigned a membership grade corresponding to the fuzzy region/fuzzy set, $F = \{low, medium, high\}$, using the membership function, $\mu_F(\underline{x}_i[k])$, where $\mu_F(\underline{x}_i[k]) \in [0, 1]$. In this work, $\mu_F(.)$, is defined as $\pi(\underline{x}_i[k], c, \lambda)$, and given by:

$$\pi(\underline{x}_i[k], c, \lambda) = \begin{cases} 2(1 - \|\underline{x}_i[k] - c\|/\lambda); \text{if } \lambda/2 \le \|\underline{x}_i[k] - c\| \le \lambda \\ 1 - 2(\|\underline{x}_i[k] - c\|/\lambda); \text{if } 0 \le \|\underline{x}_i[k] - c\| < \lambda/2 \\ 0; \text{otherwise} \end{cases} \qquad (1)$$

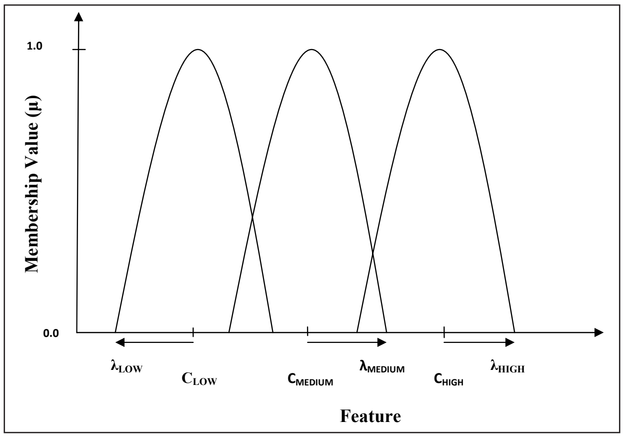
where $\lambda (> 0)$, is the scaling factor of the function $\pi(.)$, c is the central point and $\|.\|$, is the L^p-*norm*. We used the Manhattan norm in our experiments (refer Section 4).

To generate c, λ, for the fuzzy membership function, $\pi(.)$, for the fuzzy regions F, over the input range $\underline{x}_i[k]$, for $i = 1, 2, \ldots, N$ and $k = 1, 2, \ldots, p$; where N: is the number of input patterns(vectors) and p: is the features of each pattern; the following methodology is employed. Let $c_{medium(\underline{x}i[k])} = mean([\underline{x}_i[k]^-, \underline{x}_i[k]^+])$, be used to sub-divide the interval range into $[\underline{x}_i[k]^-, c_{medium(\underline{x}_i[k])})$, $(c_{medium(\underline{x}i[k])}, \underline{x}_i[k]+]$, where $\underline{x}_i[k]^-$ and $\underline{x}_i[k]^+$, are the minimum and the maximum bound of the feature $\underline{x}_i[k]$. Let $c_{low(\underline{x}_i[k])} = mean([\underline{x}_i[k]^-, c_{medium(\underline{x}_i[k])}))$ and $c_{high(\underline{x}_i[k])} = mean((c_{medium}(\underline{x}_{i[k]}), \underline{x}_i[k]^+])$. Corresponding to the three fuzzy ranges, we calculate $\lambda_{low(\underline{x}_i[k])} = \frac{1}{degOv}(c_{medium(\underline{x}_i[k])} - c_{low(\underline{x}_i[k])})$, $\lambda_{medium(\underline{x}_i[k])} = \frac{1}{2}(c_{high(\underline{x}_i[k])} - c_{low(x_i[k])})$, and $\lambda_{high(\underline{x}_i[k])} = \frac{1}{degOv}(c_{high(\underline{x}_i[k])} - c_{medium(\underline{x}_i[k])})$; where *degOv* is a parameter that controls the extent of overlap of FMF's of the three fuzzy sets (by default $degOv = 1$). These quantities are then used in Eq. (1), to fuzzify the feature space, as shown in Fig. 1.

2.2. Rough Set Theory

We will discuss some essential rough set theoretic concepts [8], relevant to this article. Let $S = (\mathbf{U}, \mathbf{E} \cup d)$, be a *decision representative system;* where $\mathbf{U}$ is a nonempty, finite set of *objects* called *universe;* $\mathbf{A} = \{\mathbf{E} \cup d\}$, is a nonempty finite set of *primitive attributes;* and $d \notin \mathbf{E}$, $d \in \mathbf{A}$, is the decision attribute (predefined output class/human activity class). For every primitive attribute $e \in \mathbf{E}$, is a function $e : \mathbf{U} \to V_e$, where V_e, is the set of values of e. With every subset of attributes BE and a subset of objects XcU, there can be associated a binary relation R_B, called *indiscernibility relation* or *equivalence relation* over $\mathbf{U}$; defined as:

Figure 1: π-Membership Function (μ) for a Feature Corresponding to Fuzzy Sets



$$\mathbf{U}/R_B = \{(x,y) \in \mathbf{X}^2 : \forall e \in B, e(x) = e(y)\} \qquad (2)$$

In Section 2.1.1, it was discussed that the set $\mathbf{X}$, although can be crisp; can be fuzzified over its feature space. With the set $\mathbf{X}$ and an equivalence relation R_b, there can be associated two subsets; called the *B-upper* and *B-lower* approximations of $\mathbf{X}$, defined are:

$$R_{B^*}\mathbf{X} = \{x \in \mathbf{U}/R_B(\mathbf{X}) : x \subseteq \mathbf{X}\} = \{x \in \mathbf{U} : [x]R_B \subseteq \mathbf{X}\} \qquad (3)$$

$$R_{B^*}\mathbf{X} = \{x\, e\, \mathbf{U}/R_B(\mathbf{X}) : x \cap \mathbf{X} \neq 0\} = \{ x \in \mathbf{U}: [x]R_B \cap \mathbf{X} \neq 0\} \quad (4)$$

The tuple $< R_{B*}, R_B{}^* >$, is called as rough set and $[x]_{R_B}$, are the equivalence classes of the equivalence relation R_B. The objects in $R_B{}^*$, signify that they can be *certainly* classified as members of $\mathbf{X}$, based on the knowledge in R_B (refer Fig 3). Similarly, objects in $R_B{}^*$, can *be possibly* classified as members of $\mathbf{X}$, based on the knowledge in R_B. In [8], alternate definitions to Eqs. (3,4), are defined as three regions of $\mathbf{X}$, called as the *positive, boundary, negative* regions and are stated as follows:

$$POS(\mathbf{X}) = R_{B*}\mathbf{X} \qquad (5)$$

$$BND(\mathbf{X}) = R_{B*}\mathbf{X} - R_B{}^*\mathbf{X} \qquad (6)$$

$$NEG(\mathbf{X}) = \{x \in \mathbf{U} : [x]_{R_B} \cap \mathbf{X} = 0\} \qquad (7)$$

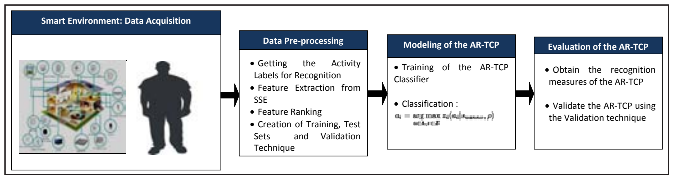
So, the Eqs. (3,4), in terms of the regions can be redefined as:

$$R_{B*}\mathbf{X} = POS(\mathbf{X}) \qquad (8)$$

$$R_B{}^*\mathrm{X} = POS(X) \cup BND(\mathbf{X}) \qquad (9)$$

Having touched on the essentials of granular computing, fuzzy sets and rough set theory; the work is well poised to assimilate the theories into adaptation of the kNN classifier.

Figure 2: Activity Recognition in a Smart Environment



3. Fuzzy-Rough Granular Computing based k-Nearest Neighbor Classifier

In the domain of activity recognition in smart environments, the user (actor) interacts with the environment to do a pre-defined set of activities. An *annotator* (observer) would collect these interactions and allocate a set of these interactions with the correct activity label (class label). These annotated collections of interactions are called as the training set. In the near future whenever the user interacts with the smart environment, the interactions are recorded and the resultant collection set is called as the test set. It is the task of the classification algorithm to provide the correct activity label to collections in the test set by using the training set, refer Fig. 2.

A *k-Nearest Neighbors* (kNN) algorithm is one of the best non-parametric classification algorithm, that assigns an activity label to an input pattern based on majority of the k-closest (using distance metrics) neighbors it finds in the training set. kNN is simple, requires no prior knowledge about the training data set, works adequately fast and has accuracy sometimes matching the ideal Bayes Classifier. kNN comes with some loopholes though; the choice of k, would either slow-down/speedup the classifier; it would always provide a classification even to absurd input patterns; when the input pattern is vague (like in the case of human activities) kNN classifier fails to capture the vagueness in the pattern; when the input pattern is incomplete (due to underlying firmware malfunctioning, or due to incomplete interactions by the user); due to such similar reasons the classifiers accuracy drops.

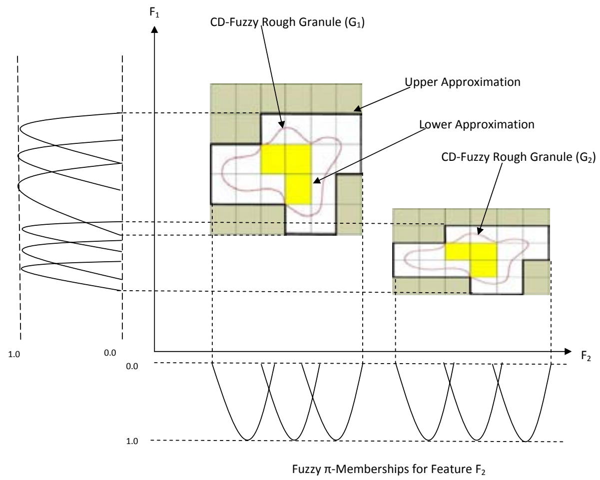
In this work, we propose a modified fuzzy-rough granular computing based kNN classifier, that hybridizes the traditional kNN, by adapting it to problems where input patterns may be vague, carry insufficient knowledge, and contain noisy/absurd, incomplete data. To address these problems, fuzzy memberships functions are learnt from the training data itself. From there, fuzzy similarity measures are defined to quantize the degree to which input patterns are similar for every feature, by granulating patterns belonging to the same decision class (these granules are called as *Class Dependent* [CD] granules). Finally, the resultant fuzzy rules are connected using various fuzzy residuum t-norms and implicators [9], to get reduced and most discriminating fuzzy rules (rules with the best classification criteria). These fuzzy rules are referred as *fuzzy equivalence relation*, that is, belongingness of every input pattern to the specified CD granule is specified by a membership function, as shown in Fig (3). Using the *fuzzy equivalence relation*, the fuzzy lower and upper approximation is calculated. These approximations are then used by the modified kNN classifier for classification.

3.1. Theory on Fuzzy Similarity and Rules Reductions via t-norm, Implicator

From Section 2.1, the training data's feature space would be fuzzified to corresponding fuzzy regions $F = \{low, medium, high\}$. Thus, for an input pattern $\underline{x}_i \in \mathbf{R}^p$, the fuzzified feature space would be $\underline{x}_i = [\pi(\underline{x}_i[1], c_{low}, \lambda_{low}), \pi(\underline{x}_i[1], c_{medium}, \lambda_{medium}), \pi(\underline{x}_i[1], c_{high}, \lambda_{high}), \ldots, \pi[p], c_{low}, \lambda_{low}), \pi(\underline{x}_i[p], c_{medium}, \lambda_{medium}), \pi(\underline{x}_i[p], c_{high}, \lambda_{high})]$; using Eq. (1). Traditional *fuzzy kNN* classifier [10], would generally take the test object $\underline{t} \in \mathbf{R}^p$ and find the FMF to k nearest neighbors training data (say $\mathbf{N}$, be set of k nearest neighbors), using the formula (10)[1]:

$$Z_{\underline{t}}(j) = \frac{\sum_{n=1}^{p} \underline{x}_i[n](1/L^p(\underline{t}(n), \underline{x}_i[n])}{\sum_{n=1}^{p}(1/L^p(\underline{t}(n), \underline{x}_i[n])} \tag{10}$$

Figure 3: Class Dependent (CD) Fuzzy-Rough Granules for Features *F1*, *F2*



The fuzzy rule to get the correct class (activity) label would simply be: $Class = \text{argmax}_{\forall j \in \mathbf{N}} Z_{\underline{t}}(j)$

We have modified the conventional *Fuzzy kNN* classifier, by partitioning the input feature space to determine the fuzzy equivalence classes. We have restricted partitioning by granulating neighbors $\underline{n} \in a$, where $a \in \mathbf{A}$. Say, $\mathbf{A}$ is the set of decision classes (human activity/concept). Such granulation of returned neighbors $\underline{n}$, over all existing decision classes $\forall a \in \mathbf{A}$, is called as *class dependent granulation*. These fuzzy equivalence classes are determined using fuzzy similarity relations for every feature between the *test object* $\underline{t}$ and *nearest neighbors* $\underline{n}$, that are returned by the traditional kNN classifier. There are various fuzzy similarity relations R, over features $e \in \mathbf{E}$, and restricted by the class dependent granule $\underline{x} = \cup\{d(\underline{n}) \in a\}, \forall_{a\in}\mathbf{A}$, where $d(.)$ returns the corresponding decision concept/class; and are given below.

$$\mu_{R_e}(\underline{x}, \underline{t}) = 1 - \frac{|e(\underline{x}) - e(\underline{t})|}{|e_{max} - e_{min}|} \tag{11}$$

$$\mu_{R_e}(\underline{x}, \underline{t}) = exp\left(\frac{-(e(\underline{x}) - e(\underline{t}))^2}{2\sigma_e^2}\right) \tag{12}$$

$$\mu_{R_e}(\underline{x}, \underline{t}) = max\left(min\left(\frac{(e(t) - e(x) + \sigma_e)}{\sigma_e}, \frac{(e(x) - e(t) + \sigma_e)}{\sigma_e}\right), 0\right)$$

[1] $L^p(.)$: Is the conventional L^p norm

In Section 4, we have experimented over few of these relations.

To consolidate the resulting FMF's using either Eqs. (11,12,13), over the subset of features w c E, the following induction is used:

$$\mu_{R_w}(\underline{x},\underline{t}) = \bigcap_{e \in w} \mu_{R_e(\underline{x},\underline{t})} \tag{14}$$

Here Eq. (14), would give the induced similarity relation over all subset of features. This is similar to getting the *antecedent* part of the fuzzy rule. To get the *consequence* part of the fuzzy rule, we introduce a novelty here. Let consequence part [i.e. the decision class(s)/activity label(s)], of the fuzzy rule be itself a fuzzy concept, $d \subseteq$ A. We propose a T-transitive fuzzy similarity relation for a fuzzy concept *d*, for a given test object $\underline{x}$, using fuzzy implicator as:

$$\mu_{R_w d}(\underline{x}) = \inf_{\underline{y} \in \mathbb{U}} \mathbb{I}\left(\mu_{R_w}(\underline{x},\underline{y}), \mu_d(\underline{y})\right) \tag{15}$$

where **I**, is the fuzzy implicator and are given in many ways:

$$\textbf{Largest } \mathbb{I}_l(x,y) = \begin{cases} 0; \text{if } x = 1 \text{ and } y = 0 \\ 1; \text{ otherwise} \end{cases} \tag{16}$$

$$\textbf{Łukasiewicz } \mathbb{I}_Ł(x,y) = min(1, 1 - x + y) \tag{17}$$

$$\textbf{Gödel } \mathbb{I}_G(x,y) = \begin{cases} y; \text{if } x > y \\ 1; \text{if } x \leq y \end{cases} \tag{18}$$

$$\textbf{Weber } \mathbb{I}_W(x,y) = \begin{cases} y; \text{if } x = 1 \\ 1; \text{if } x < y \end{cases} \tag{19}$$

Similarly, a T-transitive fuzzy similarity relation for a fuzzy concept *d*, for a given test object $\underline{x}$, using fuzzy t-norm as:

$$\mu_{R_w d}(\underline{x}) = \sup_{\underline{y} \in \mathbb{U}} \mathbb{T}\left(\mu_{R_w}(\underline{x},\underline{y}), \mu_d(\underline{y})\right) \tag{20}$$

where **T**, is the fuzzy t-norm and are given in many ways:

$$\textbf{Minimum } \mathbf{T}_M\ (x,y) = min\{x,y\} \tag{21}$$

$\mathbf{T}_M$ is also called as **Godel t-norm.**

$$\textbf{Łukasiewicz } \mathbb{T}_Ł(x,y) = max(0, x + y - 1) \tag{22}$$

$$\textbf{Weber } \mathbb{T}_{W_p}(x,y) = max\left(0, \frac{x + y + pxy - 1}{1 + p}\right);\ p \in [-1, \infty) \tag{23}$$

So having got the granulated fuzzy set $\underline{x}$ and the fuzzified test object $\underline{t}$, fuzzy similarity relation $\forall_{e \in w} R_e$, induced the subset of features $w \in \mathbf{E}$, approximation of *d* is given by:

$$\mu_{R_{e*}d}(\underline{t}) = \inf_{\underline{y} \in \underline{x}} \mathbb{I}\left(\mu_{R_e}(\underline{t},\underline{y}), \mu_d(\underline{y})\right) \tag{24}$$

$$\mu_{R_e{}^*d}(\underline{t}) = \sup_{\underline{y} \in \underline{x}} \mathbb{T}\left(\mu_{R_e}(\underline{t},\underline{y}), \mu_d(\underline{y})\right) \tag{25}$$

Finally, to allocate the correct decision class (fuzzy concept) *d*, to the test object $\underline{t}$, could be achieved by proposing a function *MaxPOS -MinBND(decision-class, testsample)*, which is defined below. This function would try to allocate the *testsample* to that *decision-class* for which the boundary region would be minimum and the positive region would be maximum:

$$MaxPOS\_MinBND(d,\underline{t}) = (w_1.\mu_{R_{e*}d}(\underline{t})\ +\ w_2.\mu_{R_e{}^*d}(\underline{t})) \tag{26}$$

$$Class = \arg\max_{d \subseteq \mathbb{A}}\ (MaxPOS\_MinBND(d,\underline{t})) \tag{27}$$

; where w_1,w_2: are the weights allocated to the lower, upper approximation. The default weights are $w_1 = 0.5$ and $w_2 = 0.5$. Recollecting rough set theory [8] and sub-Section 2.2, if $\mu R_e{}^*d(\underline{t})$ is high, it means that *all* the neighbors of $\underline{t}$ surely belong to the decision class (concept) *d;* and a high value of $\mu R_e{}^*d(\underline{t})$, means that there would be *at least one* neighbor of $\underline{t}$ belongs to decision class *d*. In [11], gives theoretical proof to support that the fuzzy lower and upper approximation in a A-class classification problem behaves in a possibilistic way; which is true in Eq. (27). Referring to Eq. (8), the fuzzy lower approximation has information regarding the magnitude of certainty of the test object to a neighbors decision class (concept). Contrarily, the upper approximation (refer Eq. (9)) holds information regarding the magnitude of uncertainty of the test object to a neighbors decision class; this information is very critical to separate between decision classes of neighbors. Say, there are two neighbors to a test object that belonging to two different decision classes; they may result in the same lower approximation, but the first neighbor may have a smaller upper approximation. Thus the neighbor having smaller upper approximation, can be translated into saying that the boundary region of that neighbor is smaller (refer Eq.(6)). According to Eq. (9), the first neighbor has less uncertainty thus it accords higher belongingness of the test object to the decision class of the first neighbor. So the largest value of the weighted sums of the fuzzy lower and upper approximations in Eq. (27), would endorse for belongingness of the test object to that decision class.

Algorithm 1 : Granular Fuzzy-Rough kNN (GrFRNN)

Input: The annotated training set $\mathbb{L} = \{l_1, l_2, ..., l_N\}$, where $l_i \in \mathbb{R}^p$.
K : Number of Nearest Neighbors
$\mathbb{A}$: Set of decision Classes, $a \in \mathbb{A}$. Such that $l_i \rightarrow \mathbb{A}_j$.
$\mathbb{E}$: Feature Set, $e \in \mathbb{E}$
$\underline{t}$: Test pattern without decision Class
Output: Allocate decision Class $\mathbb{A}_j$ for $\underline{t}$.
BEGIN:

1: $\mathbb{X} \leftarrow$ getNearestNeighbors($\mathbb{L}, K, \underline{t}$) ▹ Traditional kNN Algorithm
2: **for each** $\underline{x_i} \subseteq \mathbb{X}, \underline{x_i} \in \mathbb{R}^p$ **do**
3: $\underline{x}_i = [\pi(\underline{x}_i[1], c_{low}, \lambda_{low}), ..., ...]$ ▹ $F = \{low, medium, high\}$ ▹ Gives Fuzzified Regions
4: $\mathbb{Z}_d \leftarrow \cup\{d(\underline{x}_i) \in a\}, \forall a \in \mathbb{A}$ ▹ Class Dependent (CD) Granules
5: **end for**
6: **for each** $\underline{y} \in \mathbb{Z}_d$ in granulated fuzzified feature space $\underline{y}$ over F **do**
7: $\mu_{R_e}(\underline{y}, \underline{t}) = 1 - \frac{|e(\underline{y}) - e(\underline{t})|}{|e_{max} - e_{min}|}$ ▹ Fuzzy Similarity Relation R, $e \in \mathbb{E}$
8: $\mu_{R_{e*}d}(\underline{t}) = \inf_{\underline{y} \in \mathbb{Z}_d} \mathbb{I}\left(\mu_{R_e}(\underline{t}, \underline{y}), \mu_d(\underline{y})\right)$ ▹ Fuzzy Lower Approximation for Class d
9: $\mu_{R_e^*d}(\underline{t}) = \sup_{\underline{y} \in \mathbb{Z}_d} \mathbb{T}\left(\mu_{R_e}(\underline{t}, \underline{y}), \mu_d(\underline{y})\right)$ ▹ Fuzzy Upper Approximation for Class d
10: $MaxPOS\_MinBND(d, \underline{t}) = (w_1.\mu_{R_{e*}d}(\underline{t}) + w_2.\mu_{R_e^*d}(\underline{t}))$ ▹ Calculates MAX positive and MIN boundary regions
11: **end for**
12: **Output** $Class = \arg\max_{d \subseteq \mathbb{A}} (MaxPOS\_MinBND(d, \underline{t}))$ ▹ Allocate decision class to test pattern t

END

3.2. Proposed Algorithm of the Improvized Fuzzy-Rough Granular Computing based kNN by Learning from Data (GrFRNN)

There were numerous efforts to apply hybridized fuzzy rough concepts to kNN [10][12][11][13], to various applications. Specially the work of [13], uses fuzzy lower and upper approximations to classify test objects. In all these preceding works and in our literature survey, we have not come across work using granulating fuzzy lower and upper approximations based on class dependency of the neighbors. So also we have not come across work that fuzzifies feature space based on the linguistic {*low, medium, high*} fuzzy sets using the π-fuzzy membership function, in the domain of human activity recognition. The novelty of this work is in fuzzifying feature space using training data and in granulating nearest neighbours based on their classes belongingness. We have proposed Algorithm 1, which takes as its inputs; training set **L**, of human interactions in smart environment which are correctly labeled from set **A** by the *annotator;* and, a test pattern $\underline{t}$, for which the decision class is to be provided. At line (1), the function **X** $\leftarrow$ getNearestNeighbors (L, K, $\underline{t}$), makes a call to the traditional nearest neighbor algorithm, which returns K closest neighbors to test pattern $\underline{t}$, using some distance metric. The novelty of Algorithm 1, is in fuzzifying input feature space by using input data itself, as depicted at line (3). A novelty in the algorithm is in creating class dependent granules, by aggregating all those neighbors $\underline{x}_i$ belonging to a decision class a $\in$ **A**, by executing the function $d(\underline{x}_i) \in a$, as in lines (3,4). This eliminated the need of a domain expert in proposing FMF for the problem as discussed in Section 2.1. Using class dependent granulation of the feature space, the FMF's generated would be naturally oriented, characterizing human thinking and reasoning as discussed in Section 2. Another novelty comes with the fuzzy similarity relation R, between the nearest neighbors $\underline{x}$ got from line (1), and the test patten $\underline{t}$, at line (7). This is innovative because many of the human activities are vague (no precise definition), overlapping, having absurd/missing data items, incomplete, concurrent etc. (discussed in Section 1 and Section 2.1.1), and as such are aptly captured by the fuzzy similarity relation. There are other fuzzy similarity relations in Eqs. (12), (13); but we have used Eq. (11) for our experiments. At lines (8,9), the *consequence* part of the Fuzzy Rule is constructed for every decision class C; via the fuzzy T-transitive upper and lower approximation relation between the test pattern $\underline{t}$, and neighbors belonging to the considered decision class ($\underline{y} \in Z_d$), that was previously found using the induced fuzzy similarity relation at line (7).

From among the many options for the choice of fuzzy implicator's and fuzzy t-norm's we have used Łukasiewicz $I_Ł$ and $T_Ł$ for our experiments. Line (10), calculates the maximum positive region and the minimum boundary region for the test pattern $\underline{t}$ to the corresponding decision class. Line (12), finds the decision class to be allocated to the test pattern $\underline{t}$. Algorithm 1, has quadratic time complexity, $O(Np)$, where N: number of input pattern and p: features of each pattern.

4. Experimental Results and Discussion

This research work is biased towards the application domain of recognising human activities of the elderly population effected by dementia. Such datasets are rare and expensive to procure and produce. A publicly available Cairo dataset, from the Washington State University's CASAS testbed [14], was used for testing and validating our proposed model. This dataset has two residents and a pet, it has 27 motion sensors (other sensors are ignored), 600 activities that were recorded for a period of 3 months. These two residents are seniors with health concerns like dementia. They get occasional visitors to their homes; the visitors interactions with the smart environment is also recorded. The reasons for choosing this dataset is spelt in [3], prime among them being: (i) Non-Intrusive: The sensors used for the dataset are not audio/video based, as such they do not infringe the privacy of the individual. (ii) Non Body-worn Sensors: The sensors used for the dataset were placed scientifically in the environment and are not required to be body-worn by the individual. The dataset has the following 10 self-explaining macro-activity labels A, that were used for annotating the datasets by the annotator; Breakfast [47], Lunch [37], Laundry [10], Dinner [42], Leave home (L Home) [70], Taking medicine (R Med) [45], Night wandering (N Wand) [67], Go to Work (C work) [40], Bed [212], Bed to toilet (B Toilet) [30]. The figure given along side the macro-activity labels are the instances of that activity in the dataset and reflect the highly imbalanced nature of the dataset. This dataset was selected since activities are vague like Breakfast, Lunch, Dinner; since in a smart environment all these are eating related activities; occurring around the dining room. From the linguistic perspective we know that these activities happen at different times in the day, but at exactly what time would they occur is an issue. These vague concepts are aptly captured by relying on fuzzy-rough set theoretic concepts. This dataset has multiple residents, thus interactions would be concurrent, overlapping, sequential, sometimes incomplete; which would have to be correctly classified by the proposed Algorithm 1.

A k-fold cross validation was used, with k = 7, to validate the accuracy of the proposed Algorithm 1. The choice of k = 7, was driven due to the highly imbalanced dataset. The fuzzy similarity relation Eq. (13), was used for the setup. The Łukasiewicz Fuzzy Implicator Eq. (17) and Łukasiewicz Fuzzy t-norm Eq. (22) are used in calculating the fuzzy lower and upper approximations. One underlying motivation for this work was to use minimum, intuitive and naturally occurring feature set for the experimentation. As a result, only two feature sets were used, SENSOR[27]: corresponding to the 27 motion sensors and TOD[7]: where the time in a day was discretized to seven bins {[0-5),[5-10),[10-12),[12-15),[15-18),[18-21),[21-0)}. The K of the kNN Algorithm when set to K = 11, was seen to be best performing, for the Cairo dataset. The K of the kNN Algorithm when set to K = 11, was experimentally verified to be best performing, for the Cairo dataset. The experimental results for other values of K (neighbors), is not shown, due to space constraints. The accuracy of the proposed Algorithm 1 (GrFRNN) was compared to other classifiers like Naive Bayesian(NB), Improved Naive Bayesian (INB)[15], Traditional kNN (kNN), Fuzzy kNN by Learning from Data (FNN-LD)[16] and the result is tabulated in Table 1. The Improved Naive Bayesian (INB)[15], is a heuristic way of improving Naive Bayes', skewed data biases, systematic and weighted magnitude errors. The Fuzzy kNN by Learning from Data (FNN-LD)[16], is an improved fuzzy kNN classifier that builds its FMF's from nearest neighbors of the test object. The GrFRNN overall accuracy was the highest at 87.88%; outperforming the traditional kNN by over 16.43%, INB by over 13.03%, FNN-LD by 10.25% and NB by 28.40%.

5. Conclusion and Future Direction

The efficiency of the GrFRNN algorithm is shown in Table 1. Fig. 4, shows a graphical comparison of the accuracy of the GrFRNN versus other classifiers over different validation rounds. Figure 5, shows the box-plot of the accuracy of the GrFRNN versus other classifiers over different validation rounds. Figure 5, shows that the deviation of the GrFRNN is the minimum among its contemporaries and as such the GrFRNN is the more stable classifier. The proposed GrFRNN Algorithm,

Figure 4: Accuracy of the Various Classifier for Different Validation Rounds

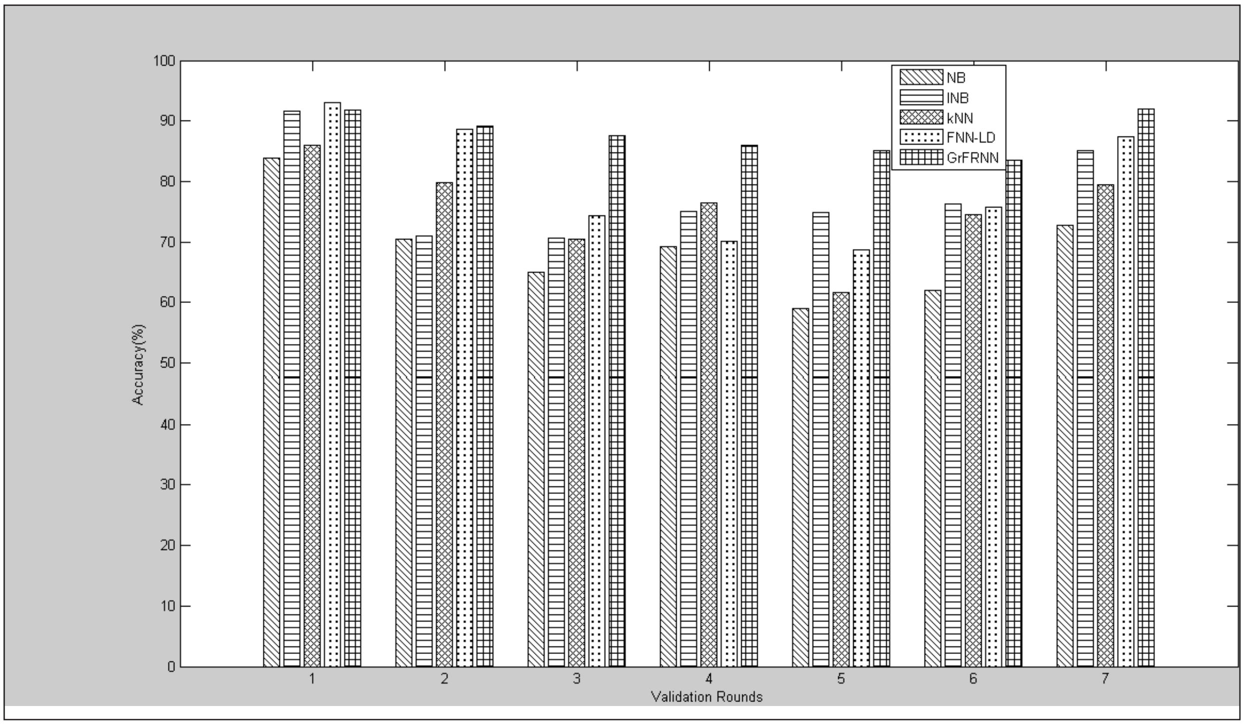


Figure 5: Box Plot of the Accuracies of the Various Classifiers for Different Validation Rounds

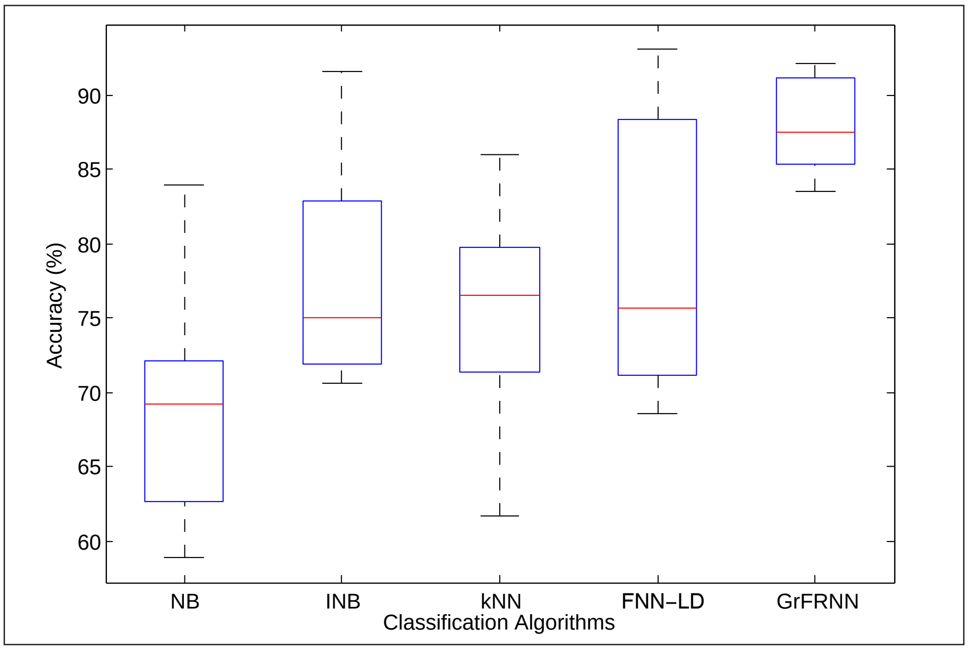


Table 1: Accuracy of the Proposed GrFRNN Algorithm Versus the Naive Bayesian, Improved Naive Bayesian, Traditional kNN, FNN-LD Approaches for Different Validation Rounds

Validation#	#Test Object	Acc. NB(%)	Acc. INB(%)	Acc. kNN(%)	Acc. FNN-LD(%)	Acc.GrFRNN (%)
1	86	83.95	91.55	85.98	93.03	91.86
2	92	70.42	70.93	79.78	88.68	89.13
3	88	64.91	70.58	70.40	74.32	87.50
4	93	69.23	75.03	76.54	70.18	86.02
5	87	58.96	74.88	61.72	68.63	85.06
6	91	61.91	76.20	74.44	75.69	83.52
7	63	72.73	85.08	79.48	87.42	92.06
Avg. (%)	100	68.44	77.75	75.48	79.71	87.88

fuzzifies the input feature space by learning from data. This is innovative as it eliminates the need of a domain expert in coming with discriminating fuzzy membership functions. Using these fuzzified regions, neighbors are aggregated together based on their class labels, to create lower and upper approximated fuzzy granules. This granulated feature space is experimentally shown to improve the accuracy of the hybridized kNN classifier in the human activity recognition domain. The proposed GrFRNN algorithm was tried on a smart environment dataset that had vague, overlapping, incomplete, concurrent human activities, which had to be properly classified. The GrFRNN Algorithm gave an improved accuracy of 87.88% and as such showed its robustness. It is observed from the Algorithm 1, that for every round of validation, it gets called afresh. As such the fuzzy lower/upper approximations, μ_{Re*d} and μ_{Re*d} at step (8,9), get re-calculated from scratch. This is reflected in the validation rounds in Table 1, where the accuracy of the GrFRNN varies from 91.86% to 83.52%. In the future, we intend to update the fuzzy lower/upper approximations, μ_{Re*d} and μ_{Re*d}, across validation rounds, instead of re-calculating from scratch. To achieve that there is work in Case-based reasoning (CBR)[17], [18] [19] that can be adopted herein. CBR is also called as transfer learning in literature, and is useful in rapid adapting of activity recognition in smart environment by limiting time spent on training of the classifier (Fig. 2).

A limitation of this work is that we have not shown experimental results of missing and absurd data on the efficiency of the GrFRNN. Another important research requirement of human activity recognition is in outliers detection, this is not touched in this work. These problems would be taken up in the future.

References

Borges, V., & Jeberson, W. (2014). Survey of Context Information Fusion for Sensor Networks Based Ubiquitous Systems, Computer Science and Information Technology 2(3) 165–178. doi:10.13189/csit.2014.020306. URL http://www.hrpub.org/journals/jour archive.php?id=35

Wimo, A., Winblad, B., Jo¨nsson, L. (2010). The worldwide societal costs of dementia: Estimates for 2009, Alzheimer's & Dementia 6(2), 98-103.

Borges, V., & Jeberson, W. Modeling Activity Recognition Using Text Categorization Paradigm, Elsevier Expert System with Applications (under consideration).

Bargiela, A., & Pedrycz, W. (2003). Granular computing: An introduction, Springer.

Pedrycz, W., Skowron, A., & Kreinovich, V. (2008). *Handbook of granular computing*. John Wiley & Sons.

Pal, S. K., & Dutta-Majumder, D. K. (1986). *Fuzzy mathematical approach to pattern recognition*. Halsted Press.

Bottou, L., & Vapnik, V. (1992). *Local learning algorithms*. *Neural Computation*, 4(6), 888-900.Z. Pawlak, Imprecise Categories, Approximations and Rough Sets, Springer, 1991.

Pawlak, Z. (1991). *Imprecise categories*. *Approximations and rough sets*. Springer.

Mundici, D. (2011). *Advanced Lukasiewicz calculus and MV-algebras*. Springer.

Keller, J. M., Gray, M. R., & Givens, J. A. (1985). *A fuzzy k-nearest neighbor algorithm*. IEEE Transactions on Systems, Man and Cybernetics, (4), 580-585.

Sarkar, M. (2007). *Fuzzy-rough nearest neighbors algorithm*. Fuzzy Sets and Systems, 158, 2123-2152.

Shen, H. B., Yang, Y., Chou, K. C. (2006). *Fuzzy KNN for predicting membrane protein types from pseudo-amino acid composition*. Journal of Theoretical Biology, 240(1), 9-13.

Jensen, R., & Cornelis, C. (2011). *Fuzzy-rough nearest neighbour classification and prediction*. Theoretical Computer Science, 5871-5884.

WSU, CASAS Project (2014). Retrieved from http://ailab.wsu.edu/casas/

Borges, V., & Jeberson, W. Fortune at the bottom of the Classifier Pyramid: A Novel approach to Human Activity Recognition. Elsevier Procedia Computer Science.

Borges, V., & Jeberson, W. (2014). *Modification of fuzzy kNN by learning from data in activity recognition modeling*. International Conference on Mathematics and Computational Sciences (In Press), Inderscience Publishers.

Pal, S. K., & Mitra, P. (2004). *Case generation using rough sets with fuzzy representation*. IEEE Transactions on Knowledge and Data Engineering, 16(3), 293-300.

Li, Y., Shiu, S. C. K., & Pal, S. K. (2006). *Combining feature reduction and case selection in building CBR classifiers*. IEEE Transactions on Knowledge and Data Engineering, 18(3), 415-429.

Li, Y., Shiu, S. C. K., Pal, S. K., & Liu, J. N. K. (2006). *A rough set-based case-based reasoner for text categorization*. International Journal of Approximate Reasoning, 41(2), 229-255.